

Lack of Ability is Not Always the Reason for High Attrition¹

Thomas R. Carretta Malcolm James Ree

Air Force Research Laboratory
Human Effectiveness Directorate
Brooks Air Force Base, TX

Introduction

Training employees is integral to many organizations. When training attrition occurs, it is not unusual for managers to ask human resources specialists to conduct validation studies and set higher standards for selection. This may help in many cases, but sometimes is not fruitful. The following is an example of a training attrition situation where a valid predictor exists, is in place, but cannot be expected to reduce attrition. Other factors such as motivation and job design must be considered.

The job of weapons director (WD), once only for commissioned officers, has recently become available to enlisted U. S. Air Force (USAF) members. Currently, WD teams can be made up of both officers and enlisted members. The goal is to have one officer as a supervisor for about five enlisted members.

All enlisted members must be in their second four-year term of enlistment. This means that they have already been trained in another USAF technical specialty. Soon after enlisted WD training began, attrition problems were observed. Across the USAF typical attrition in enlisted technical training is about 5%. In the enlisted WD training it has run about 9% for several years, but 20% for the most recent year (1996). Because selection was by appointment or voluntary and because no test score minimums existed, senior managers sought a solution in the setting of standards for a valid selection composite. This effort is reported in the current study.

Method

Participants

The participants were 353 enlisted Air Force members who attended WD training. All were in their second enlistment tour. They had successfully completed technical training for and performed on the job in their prior technical specialty. Of these, 321

¹ Previously published as Carretta, T. R., & Ree, M. J. (1997). Lack of ability is not always the reason for high attrition. *Proceedings of the 39th Annual Conference of the International Military Testing Association*, Sydney, Australia, 71-75.

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 1997		2. REPORT TYPE		3. DATES COVERED -	
4. TITLE AND SUBTITLE Lack of Ability is Not Always the Reason for High Attrition				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Research Laboratory, Human Effectiveness Directorate, Brooks AFB, TX, 78235				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT see report					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 8	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

successfully completed WD training and 32 attrited from training. All enlisted WD data were collected between 1990 and 1996.

They were selected for USAF enlistment, at least in part, by scores on the Armed Services Vocational Aptitude Battery (ASVAB; Ree & Carretta, 1994). Due to this selection, they constituted a range restricted sample. There is no official screening process for enlisted WD training and almost all trainees were non-volunteers.

Measures

Predictors. The ASVAB is comprised of 10 tests that measure general cognitive ability (g) and the lower-order factors of verbal/math, speed, and technical knowledge (Ree & Carretta, 1994). The 10 tests are General Science (GS), Arithmetic Reasoning (AR), Word Knowledge (WK), Paragraph Comprehension (PC), Numerical Operations (NO), Coding Speed (CS), Auto and Shop Information (A/S), Mathematical Knowledge (MK), Mechanical Comprehension (MC), and Electronics Information (EI).

The predictors were composites of the ASVAB tests. The USAF, like the other services, aggregates the tests into composites to create more reliable scores and to have a broad content for the composites. The Air Force composites include Mechanical ($M = MC + GS + 2A/S$), Administrative ($A = NO + CS + WK + PC$), General ($G = WK + PC + AR$), and Electronic ($E = AR + MK + EI + GS$). The composite scores are reported as normative percentiles.

The literature suggests that prior educational achievement predicts future educational achievement. Therefore, two other predictors available for all participants were examined. The first of these was numbers of years of (non-USAF) schooling (EdYrs). The second was final school grade (PFSG) in previous USAF technical training.

Criterion. The criterion for the regressions was the final school grade (FSG) earned at technical training. FSG was the proportion correct in a series of paper-and-pencil written tests. The maximum was 99 and the minimum passing score was 70. If a trainee earned grades below 70 they were dropped from training for academic reasons.

During training there are a series of work samples called “progress checks” that involve realistic simulation of the job tasks. These are scored pass-fail. Trainees who fail progress checks several times can be recommended for elimination by instructors for non-academic (called “lack of progress”) reasons. A review of these recommendations is conducted by senior managers who either concur or retain the trainee. The training graduates are those who acquired the job knowledge tested in academic training and acquired the relevant job knowledge and skills tested by the progress checks.

Analyses

Participants were characterized as having completed ($n = 321$) or attrited ($n = 32$) from training. Among those who attrited from training, two reasons could be determined; academic failures and other failures. Mean scores of those who completed and failed to complete training were computed for the ASVAB composites, years of education, and prior technical training final school grade.

Those participants who attrited from enlisted WD training had either no or incomplete criterion data and could not be included in regression analyses. For the participants who completed technical training, regressions of the FSG on the four ASVAB composites, years of education, and PFSG were computed. The Type I error rate was specified as $p < .01$. After statistical significance testing, these data were corrected for range restriction (Lawley, 1943) and the same regressions again computed.

To correct for the upward-bias in \underline{R} , we applied Stein's Operator (Kennedy, 1988; Stein, 1960) to estimate shrinkage in cross-application. Stein's Operator estimates the mean \underline{R} coming from all possible cross-applications of the regression coefficients in samples of the size upon which the coefficients were estimated. It takes into account the sample size, number of predictors, and the estimated multiple correlation.

Finally, minimum cut scores were established with reference to dual goals. The first goal was to minimize attrition in training and the second goal was to reduce false positives. The goal of minimizing training attrition was investigated by inspecting predicted scores in training from the regression equation and finding the value below which training grades are expected to be less than 70. Reduction of false positives was investigated by means of two-way distributions of passing/failing and a selector composite.

Results and Discussion

As would be expected for a sample of participants subjected to prior selection, the standard deviations of the composites were reduced from the normative sample. For M, A, G, and E the standard deviations of the composites were 67%, 59%, 47%, and 53% of their respective normative values. Table 1 shows the means, standard deviations, and correlations of all variables before and after correction for range restriction.

The simple bivariate correlation of each composite with the FSG criterion can be found in Table 1. All of these correlations were statistically significant. Because the corrected correlations are superior statistical estimates we shall limit our discussion to them. The G composite ($\underline{r} = .58$) was most predictive followed closely by the E composite ($\underline{r} = .56$). The M composite was less predictive ($\underline{r} = .43$) as was the A

Clearance number PA# 97-450

composite ($r = .40$). EdYrs ($r = .35$) and PFSG ($r = .49$) were also less predictive than either G or E.

Table 1. Means, Standard Deviations, and Correlations for the Variables

	M	A	G	E	EdYrs	PFSG	FSG	Mean	SD
M	1.00	.13	.43	.64	.07	.19	.15	71.52	19.35
A	.47	1.00	.43	.35	.06	.14	.08	70.88	18.60
G	.71	.77	1.00	.70	.14	.29	.31	73.47	13.59
E	.82	.70	.93	1.00	.15	.23	.25	71.66	15.18
EdYrs	.24	.25	.33	.33	1.00	.18	.23	13.02	.82
PFSG	.45	.44	.56	.54	.30	1.00	.31	87.80	6.01
FSG	.43	.40	.58	.56	.35	.49	1.00	88.81	4.72
Mean	50.33	51.41	51.00	50.33	12.80	84.61	86.36		
SD	28.81	31.36	29.19	28.83	.85	6.97	5.50		

Note. Values above the diagonal are observed data and those below have been corrected for range restriction.

FSG was regressed on each composite, EdYrs, and PFSG. Table 2 shows the multiple correlations from the regressions. Again, the E and G composites were preferable to either the M or A composites as they were more valid. Because the multiple correlation is an upwardly-biased estimator, it is necessary to reduce the \underline{R} as would occur in cross-application of the regression weights in another sample. Stein's Operator (Kennedy, 1988; Stein, 1960) was applied and the resultant shrunken \underline{R}_s are in the last column of Table 2. These values are those that would be expected when the regression equation was validated on a separate set of applicants.

Table 2. Regression Results

Predictors	R	\underline{R}_c	\underline{R}_s	Regression Equations
M, EdYrs, PFSG	.37	.57	.55	$46.44 + .04M + 1.23\text{EdYrs} + .25\text{PFSG}$
A, EdYrs, PFSG	.36	.56	.54	$45.60 + .03A + 1.24\text{EdYrs} + .27\text{PFSG}$
G, EdYrs, PFSG	.42	.63	.61	$55.72 + .07G + .95\text{EdYrs} + .17\text{PFSG}$
E, EdYrs, PFSG	.40	.61	.59	$54.08 + .07E + .97\text{EdYrs} + .19\text{PFSG}$

Note. \underline{R} is the uncorrected observed multiple correlation, \underline{R}_c is the corrected multiple correlation. \underline{R}_s is \underline{R}_c shrunken to estimate cross-validated values.

The incremental validity of adding the additional information in the EdYrs and PFSG variables was rather small. The practical effect was an approximately .03 increase in \underline{R} . Further, the choice between E and G cannot be made on predictive efficiency grounds alone. Practical considerations are also important.

The distributions of the E composite for women and men differ. At the normative 50th percentile, 59% of the men qualify but only 42% of the women qualify. For the G composite the same values are 54% and 47%, a much smaller difference. Given that the E and G composites have similar validities it would minimize adverse impact to use the G composite.

After identifying the preferred composite it is necessary to establish minimum cut scores for operational use. The cut score should address the goals of the selecting agency, such as minimizing false positives, while being mindful of the cost of recruiting. By evaluating the predicted scores from the regression equation we can estimate the expected training grade performance of applicants. For the G composite normative percentiles of 20, 50, and 80 the predicted training grades are: 83, 86, and 90. False positives were found to occur for the G composite score ranging rather uniformly from 34 to 86 with one case at 99. Fully half the failures have G composite scores of 70 or better. This suggests that attrition was not caused by a lack of cognitive ability. At this score level, 70, the attrition would be 5%, the USAF average. However, the cost of recruiting and demand for high scoring airmen is great.

We first looked at the scores of the participants who were classified as attritees from training. Of the 32, only three were academic failures. Academic failure results when a student does not achieve sufficiently high scores on the course tests. The causes of attrition for the other 29 was a failure to progress in training. These participants performed poorly on the work sample progress checks. The mean score on the G composite for the three academic failures was 55. The predicted training grade for G = 55 was about 87. These three academic failures should have completed training. Their success in prior technical training and success on the prior job also suggests that they should have completed this training.

For the 29 non-academic failures the mean G score was 67. For the graduates the mean G score was 73, not much above the non-academic attritees. The distribution of G scores for the graduates ranged from 39 to 99, almost identical (34 to 99) to the distribution of the G scores for the non-academic attritees. In fact, the distributions are very similar with only small differences in shape and cumulative proportions. There were no notable differences in ability between those who successfully completed training and those who failed to complete training for non-academic reasons.

We speculate that the non-academic attritees failed to complete training for what might be termed "lack of motivation." First, almost all were non-volunteers. This may have had a demotivating effect. Second, the job characteristics may be demotivating. For example, extensive travel up to 220 days per year is required. Further, unlike former officer weapons directors trainees who left the Air Force if they failed training, enlisted trainees can return to other military occupations. Some trainees may have deliberately performed poorly in the progress checks to be released from training. With an average G composite

of 73 for all accessions, it is clear that this job has high ability trainees. It may be that this is not an attractive job as it is currently structured and managed.

There are several corrective paths that might be explored. The utility of non-cognitive measures could be examined for predictiveness. Conscientiousness, a personality variable that addresses motivation, has been shown to be predictive of job performance across a wide variety of occupations. Barrick and Mount, (1991) conducted a meta-analysis of the predictiveness of several personality variables for many job performance criteria. These criteria included training performance, job performance, tenure and salary. Barrick and Mount estimated the mean true validity of conscientiousness as .22 with little variability due to job type such as professional, blue collar, or sales.

However, it is unlikely that the non-academic attritees lacked in conscientiousness. They have been through basic military training, Air Force technical training, and completed one four-year term of enlistment.

If some characteristics of the job were found to be demotivating, job redesign could be undertaken to change onerous aspects such as travel requirement or long hours. Campion (1988) has developed a self-report instrument for job redesign that covers four areas: Motivation, Mechanical, Biological, and Perceptual/Motor. Campion's instrument could be used to assess and recommend changes to job design.

The impact of bonuses or other special incentives could be estimated. Finally, these three changes (non-cognitive measures, job redesign, incentives) might be used in some combination.

Although managers frequently perceive that raising standards will solve training attrition problems, this is not always the case. In the situation reviewed here there was a valid predictor available, trainees were highly selected on ability (and probably conscientiousness), but attrition from training was still high. In a case such as this, raising selection standards for ability cannot be expected to reduce attrition. The problem appears to be in motivation or in job design. To provide answers for managers these other aspects, motivation and job design must be evaluated and appropriate remediation applied.

References

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.

Campion, M. A. (1988). Interdisciplinary approaches to job design: A constructive replication with extensions. *Journal of Applied Psychology*, 73, 467-481.

Kennedy, E. (1988). Estimation of the squared cross-validity coefficients in the context of best subtest regression. *Applied Psychological Measurement*, 12, 231-237.

Lawley, D. N. (1943). A note on Karl Pearson's selection formulae. *Proceedings of the Royal Society of Edinburgh*, 62 (Section A, Pt. 1), 28-30.

Ree, M. J., & Carretta, T. R. (1994). Factor analysis of the ASVAB: Confirming a Vernon-like structure. *Educational and Psychological Measurement*, 54, 459-463.

Stein, C. (1960). Multiple regression. In I. Olkin et al. (Eds.). *Contributions to probability and statistics*. Stanford, CA: Stanford University Press.